

Prefácio da segunda edição

Boas vindas à segunda edição do livro “R para Ciência de Dados”! Esta é uma revisão importante da primeira edição, removendo material que não consideramos mais útil, adicionando material que gostaríamos de ter incluído na primeira edição e atualizando o texto e o código de acordo com as melhores práticas. Também estamos muito animados em dar as boas-vindas a uma nova co-autora: Mine Çetinkaya-Rundel, uma renomada educadora em ciência de dados e uma de nossas colegas na Posit (conhecida anteriormente como RStudio).

Segue abaixo um breve resumo das maiores mudanças nesta edição:

- A primeira parte do livro foi renomeada para “Visão geral”. O objetivo desta seção é fornecer a você os detalhes aproximados da “visão geral” da ciência de dados antes de nos aprofundarmos nos detalhes.
- A segunda parte do livro é chamada “Visualizar”. Esta parte oferece uma cobertura mais abrangente das ferramentas e práticas recomendadas de visualização de dados em comparação à primeira edição. O melhor lugar para obter todos os detalhes ainda é o [livro do ggplot2](#), mas agora o R4DS aborda de forma mais abrangente as técnicas importantes.
- A terceira parte do livro agora é chamada de “Transformar” e inclui novos capítulos sobre números, vetores lógicos e valores faltantes (*missing values* - *NA*). Esses tópicos eram anteriormente parte do capítulo de transformação de dados, mas precisavam de mais espaço para cobrir todos os detalhes.
- A quarta parte do livro é chamada de “Importar”. São novos capítulos que vão além da leitura de arquivos de texto simples para trabalhar com planilhas, importar dados de bancos de dados (*databases*), trabalhar com grandes conjuntos de dados (*big data*), converter dados hierárquicos em dados tabulares e extrair dados de sites da web (*web scraping*).
- A parte “Programar” permanece, mas foi reescrita do zero para se concentrar nas partes mais importantes da escrita de funções e iteração. Agora, a escrita de funções inclui detalhes sobre como criar funções no estilo tidy (lidando com os desafios de avaliação tidy (*tidy evaluation*)), uma vez que isso se tornou muito mais fácil e importante nos últimos anos. Adicionamos um novo capítulo sobre funções importantes do *R base* que você provavelmente encontrará em códigos em R por aí.
- A parte de modelagem foi removida. Nunca tivemos espaço suficiente para fazer justiça à modelagem e agora existem recursos muito melhores disponíveis. Geralmente, recomendamos o uso do pacote [tidymodels](#) e a leitura do livro [Tidy Modeling with R](#) escrito por Max Kuhn e Julia Silge.
- A parte de “Comunicar” permanece, mas foi atualizada em detalhes para apresentar o [Quarto](#) em vez do R Markdown. Esta edição do livro foi escrita no Quarto e ele claramente é a ferramenta do futuro.